CS 175: Projects in AI

AI For Social Good

Authors: Ananya Kashyap, Bhavya Gupta, Tanush Goel, Sebu Eisaian

Hungry Hungry Hippos

March 2024

Collaborators: Professor Nadia Ahmed, Junyao Wang, Zixiao Zong, Edgar Robles, Jacob Campbell, Jasper Doan, Tyler Fermanian, Sharon Ma
(Extending Appreciation to IEEE GRSS, the 2021 Data Fusion Contest provided us our dataset)

INTRODUCTION

In an ambitious endeavor, this project seeks to harness the extensive capabilities of satellite imagery for the pivotal task of mapping non-electrified settlements across Africa. Africa stands as the most unelectrified continent on the globe, with Sub-Saharan Africa being the most unelectrified region worldwide. The initiative is aimed at narrowing the stark electrification divide, leveraging a wealth of high-resolution data from a variety of satellite sources including Sentinel-1, Sentinel-2, Landsat 8, and VIIRS.

This endeavor required an intricate data processing methodology, setting the stage for the application of sophisticated deep learning techniques. At the heart of our analytical approach was the utilization of PyTorch Lightning, a framework that streamlined the training of complex neural network models such as U-Net, Segmentation CNN, and ResNet, each carefully chosen for their capacity to decode the intricate patterns presented in geospatial data.

The primary goal of this project was to employ the nuanced analytical power of convolutional neural networks (CNNs) to delineate and pinpoint areas devoid of electricity, using satellite imagery as a foundational data source. Spanning from the meticulous preprocessing of satellite data to the development and nuanced optimization of deep learning models, our approach embraced a comprehensive scope. Paramount to our methodology was the strategic fine-tuning of model parameters, facilitated through the integration of the Weights & Biases platform, which enabled a rigorous, experimental refinement process. This extensive model training and optimization effort was propelled by the robust capabilities of a GPU-enhanced computational environment, ensuring the efficient handling of voluminous datasets and the intricate architecture of our chosen models.

The zenith of our project was marked by the formulation of a Segmentation CNN model distinguished by its high accuracy. This achievement not only underscores a significant leap in the application of deep learning for the analysis of satellite imagery but also heralds promising prospects for significantly contributing to global electrification initiatives. Through a concerted effort encompassing advanced data processing, model development, and exhaustive parameter fine-tuning, we have forged a novel path towards leveraging cutting-edge technology to shed light on some of the most isolated and underserved communities worldwide.

The objectives of this project were multifaceted, including:

- The preprocessing of high-resolution satellite imagery from a diverse array of sources to prime the data for in-depth analysis.
- The development and meticulous training of advanced deep learning models, specifically tailored to extract and interpret electrification indicators from satellite data.
- The comprehensive evaluation of model efficacy through a suite of performance metrics, including the Jaccard Index, Intersection over Union, Accuracy, AUC, and F1-score, to verify and ensure the precision and reliability of our models.

Identifying the Need

The core need for this research stems from the critical global challenge of identifying non-electrified settlements, particularly in remote and underserved regions in Africa. Electrification is a fundamental enabler of socio-economic development, providing access to education, healthcare, and economic opportunities. Historically, the lack of reliable and up-to-date information on electrification status has hindered targeted intervention efforts. This research aims to leverage advancements in satellite imagery and computer vision to address this gap, offering a scalable and accurate method for mapping electrification globally.

Why Has This Been a Challenge in the Past?

Challenges with Computer Vision:

- Early computer vision models lacked the sophistication to accurately interpret complex satellite imagery, struggling with variations in lighting, cloud cover, and seasonal changes.
- Limited computational resources restricted the ability to process large datasets, impeding model training and deployment on a global scale.

Challenges with Using Satellite Images:

- High variability in satellite data quality, resolution, and spectral bands across different sources complicated the creation of standardized datasets for analysis.
- The vast amount of data generated by satellites presented significant storage and processing challenges, necessitating substantial computational power.

Why Hasn't This Been Pursued Till Now:

- Only recently have advances in deep learning and computer vision, combined with increased access to high-performance computing resources, made it feasible to analyze satellite imagery at scale.
- There isn't a strong and direct profit incentive, the work is seen as purely altruistic, meaning less people have been incentivized to do such work.
- Previous efforts may have been limited by the availability and cost of high-resolution satellite imagery, which has become more accessible in recent years.

State of the Art

Recent satellite competitions and models have pushed the boundaries of what's possible in satellite imagery analysis. The DeepGlobe Satellite Challenge and the SpaceNet competitions have fostered innovations in building detection, land cover classification, and road extraction from satellite images. In the future better models, onset of new algorithms, quality of data, among other improvements would certainly enhance the Hungry Hungry Hippos ability to classify non-electrified settlements. These competitions highlight the growing capability of computer vision models to tackle complex spatial tasks with high accuracy.

State of Models at This Point:

- Computer vision models, particularly Convolutional Neural Networks (CNNs), have seen remarkable advancements, enabling detailed image classification, object detection, and segmentation tasks. This works very well for our problem structure as we aim to segment specific tiles of imagery and classify them to 4 potential class labels.
- Innovations in model architectures, training methodologies, and data augmentation techniques have significantly improved the accuracy and efficiency of computer vision applications.

State of the Topic the Model Addresses:

- The current understanding of satellite imagery analysis has evolved, with increased emphasis on multi-spectral and temporal data to capture a comprehensive view of the Earth's surface.
- Advancements in machine learning algorithms have enabled the extraction of actionable insights from satellite data, facilitating applications in environmental monitoring, urban planning, and disaster response.

Related Works

- Global Forest Watch (GFW) leverages high-resolution satellite imagery and computer vision to monitor deforestation and forest degradation in real-time. This initiative, documented in Hansen, M. C., et al. (2013) Science, has significantly advanced global forest monitoring by enabling precise detection of changes in forest cover.
- Radiant Earth Foundation uses machine learning on Earth observation data to support sustainable development across various sectors, including agriculture and disaster response. Highlighted in Robinson, C., et al. (2019) Remote Sensing of Environment, the foundation's efforts in land cover classification demonstrate the impact of combining satellite data with machine learning for environmental analysis.

Data Description

Data Sources: The study utilizes the IEEE GRSS 2021 Dataset, featuring 98 tiles of 800x800 pixels, each corresponding to a 64 km^2 area. The dataset is enriched with imagery from Sentinel-1, Sentinel-2, Landsat 8, and VIIRS satellites, covering diverse spectral ranges and resolutions.

Gathering and Selecting Data: Data for training and testing the model includes:

- Sentinel-1 SAR data, focusing on VV and VH polarization for texture and moisture content.
- Sentinel-2 multispectral imagery, excluding the cirrus band, for vegetation health and land cover.
- Landsat 8 data for detailed VNIR, SWIR, and TIR analysis.
- VIIRS nighttime data for identifying electrified areas.

Semantics of the Data: Each satellite contributes unique information:

- **Sentinel-1:** Offers SAR imagery for surface texture analysis.
- **Sentinel-2:** Provides multispectral data for environmental monitoring.
- **Landsat 8:** Enhances with multispectral and thermal imagery.
- **VIIRS:** Captures nighttime light, indicating electrification.

Relation to the Problem: The comprehensive dataset enables the identification of non-electrified settlements by analyzing various indicators such as infrastructure presence (daytime imagery) and artificial lighting (nighttime imagery), directly addressing the challenge of electrification in Africa.

Size of the Data: The dataset comprises:

- **Sentinel-1:** 4 images, 2.1 GB each.
- **Sentinel-2:** 4 images, 6.2 GB total.
- **Landsat 8:** 3 images, 8.5 GB total.
- **VIIRS:** 9 images, 1.2 GB total.

Semantic labels for human settlements and electrification presence are provided, facilitating classification tasks focused on identifying human settlements without electricity.

Methodology: Leveraging Satellite Imagery for Electrification Analysis

Theoretical Aspects:

The theoretical foundation of our project is anchored in the premise that satellite imagery, when processed through advanced computer vision and machine learning techniques, can reveal critical insights into the electrification status of remote and underserved areas. We hypothesize that specific patterns and indicators within these images, can be effectively identified to map non-electrified settlements alongside 3 other types of

Machine Learning/Data Mining Techniques Used:

Our methodology encompasses a series of steps designed to harness the power of satellite imagery for predictive analysis:

Data Preprocessing: Utilizing src/preprocessing/subtile_esd_hw02.py for subtiling the satellite images into smaller, manageable pieces while preserving relevant metadata for analysis.
Custom Dataset and DataLoader: Developing a PyTorch Dataset class (src/esd_data/dataset.py) and utilizing PyTorch Lightning DataModule (src/esd_data/datamodule.py) for efficient data handling, including custom transformations for data augmentation to enhance model robustness.

Specific Algorithms Developed for the Task:

Dimensionality Reduction: Applied techniques such as PCA, TSNE, and UMAP to distill the most informative features from the satellite imagery, facilitating a focused and efficient model training process.
Segmentation CNN: A straightforward convolutional neural network designed for segmentation tasks, capturing spatial hierarchies and patterns within the satellite imagery.
Transfer Learning with ResNet101: Leveraged the pre-trained FCNResnet101 model (src/models/supervised/resnet_transfer.py), adapting it for the segmentation of satellite images by modifying the input and output layers to match our dataset specifications.
U-Net Architecture: Implemented a U-Net model (src/models/supervised/unet.py) that uses skip connections to retain important spatial information at various resolutions, enhancing the model's ability to accurately segment electrified vs. non-electrified areas.

Training and Validation:

- **Model Training:** Utilized PyTorch Lightning for structuring training and validation loops, simplifying the code and improving readability and maintenance. Training was

monitored and optimized using Weights & Biases, allowing for extensive hyperparameter tuning and performance tracking.
- **Hyperparameter Tuning:** Conducted systematic sweeps (configured in scripts/sweeps.yml) to identify the optimal settings for learning rate, batch size, and architectural parameters, ensuring the best possible model performance.

Evaluation Metrics:

- Models were evaluated based on a variety of metrics including Jaccard Index, Intersection over Union (IoU), Accuracy, AUC, and F1-score, providing a comprehensive understanding of model performance across different dimensions of accuracy and reliability.

This methodical approach, grounded in solid theoretical understanding and employing cutting-edge machine learning techniques, has enabled us to develop specific algorithms tailored to the task of identifying non-electrified settlements through satellite imagery. Through rigorous training, validation, and evaluation, we have created a highly accurate model capable of contributing meaningful insights toward global electrification efforts.

Replicability

Ensuring the replicability of our project is paramount to its success and contribution to the field. To facilitate this, comprehensive documentation is provided in the README of our repository, located at https://github.com/cs175cv-w2024/final-project-hungry-hungry-hippo. This README offers detailed instructions on the initial steps required to replicate our study, including data retrieval and setup processes.

Retrieving the Dataset:

- Our project utilizes the dfc2021_dse_train.zip dataset, which is a crucial component of our analysis. We have simplified the data acquisition process by providing a direct download link in our README, eliminating the need for manual registration with the IEEE DataPort.
- Once downloaded, the dataset should be unzipped and the Train directory placed into the data/raw directory of the project structure. This organization ensures that subsequent data processing and model training scripts function correctly, adhering to the expected file paths.

Tools, Compute, and Software Environment

To accomplish the objectives of this project, a specialized set of tools, computing resources, and software environments were meticulously selected and utilized. Below is a detailed overview:

Suggested Hardware:

- **GPU:** High-performance GPU computing was essential for the training of deep learning models. Nvidia GPUs, such as the GeForce RTX 4060 series utilized by the Hungry Hungry Hippos, were utilized to accelerate the training process, allowing for rapid iteration and experimentation. The GPU's computational power significantly reduced model training time, enabling the handling of large satellite image datasets and complex neural network architectures.

Frameworks and Software Environment:

PyTorch: Served as the primary deep learning framework due to its flexibility, efficiency, and user-friendly interface. PyTorch's dynamic computation graph enabled intuitive model development and debugging.
PyTorch Lightning: Leveraged to further streamline the training process, PyTorch Lightning abstracted much of the boilerplate training code, allowing for a more organized and readable codebase. It facilitated the use of advanced training techniques such as mixed-precision training and multi-GPU training.
Weights & Biases (W&B): Integrated for experiment tracking and hyperparameter tuning. W&B provided a comprehensive platform for logging experiments, visualizing results, and comparing different model versions, which was crucial for the iterative process of model optimization.
OpenCV and PIL: Utilized for image processing tasks, including reading, writing, and transforming satellite images into formats suitable for model training.
scikit-learn: Employed for preprocessing and dimensionality reduction tasks, aiding in feature extraction and analysis before feeding data into the neural networks.
NumPy and Matplotlib: Essential for data manipulation and visualization, enabling the analysis of satellite data and the interpretation of model outcomes.
CUDA: Crucial for efficient tensor operations and key computational tasks during model training.

Tests

The testing phase of our project was an extensive and critical component, ensuring the robustness and effectiveness of our models. This phase encompassed a rigorous schedule of training, parameter optimization, and model evaluation, which unfolded over approximately 100 hours of computational time. The testing was conducted on both GPU and CPU platforms to balance between computational efficiency and resource availability.

Training Duration and Computational Resources:

- Our models underwent an intensive training regimen, with some processes allocated to high-performance GPU computing for speed and efficiency, while others utilized CPU resources. This approach allowed us to maximize our computational budget and time, despite the limitations imposed by hardware availability.

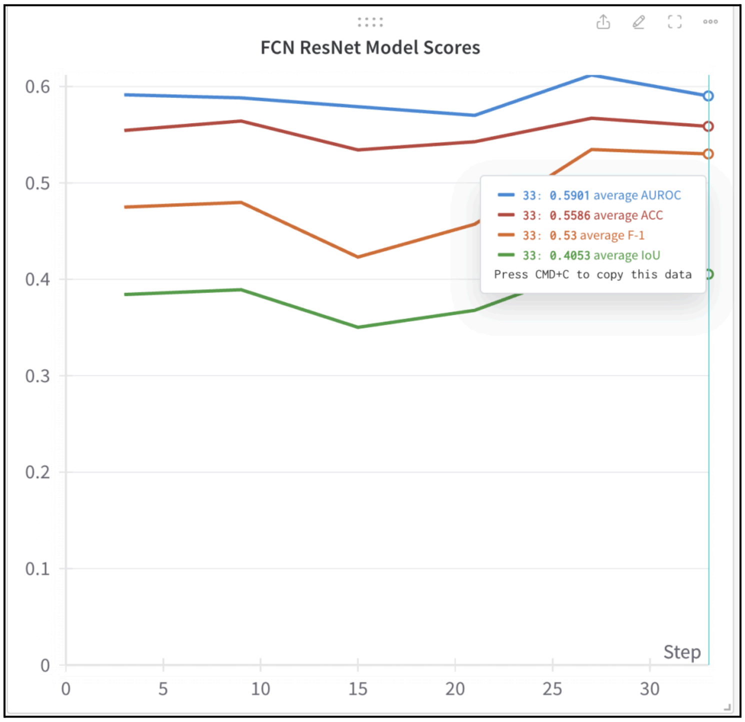
Hyperparameter Sweeps:

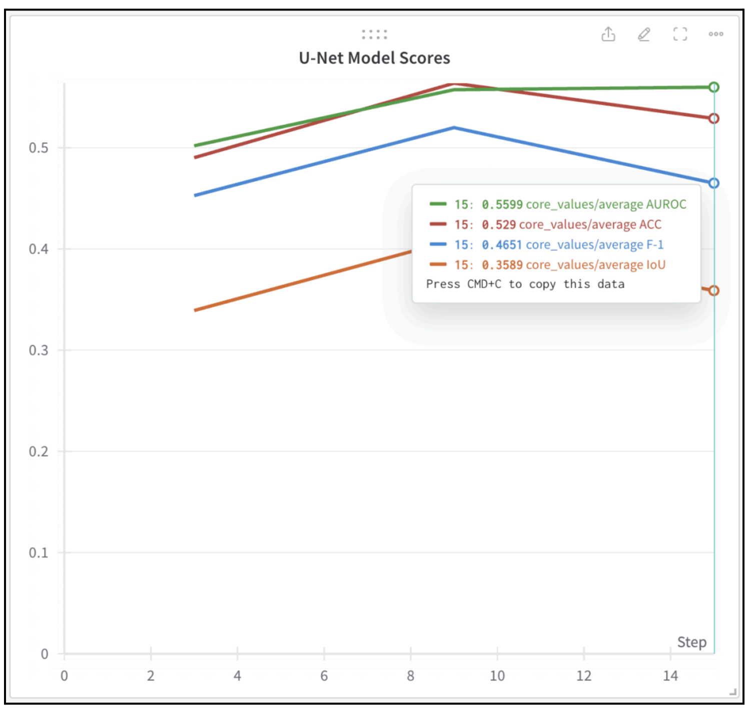
- A substantial portion of the testing phase was dedicated to hyperparameter optimization. Through extensive sweeps, we explored a range of values for key parameters to identify the combinations that yielded the best performance for each model.
  ResNet Adaptations:
  - For the ResNet model, our primary focus was on adjusting the learning rate, maximum epochs, and batch size. These parameters were meticulously tested to fine-tune the model's learning process and adapt it effectively to our dataset.
- Segmentation CNN (SegCNN) Adjustments:
  - The testing for SegCNN extended beyond the basic parameters to include the network's depth. In addition to learning rate, maximum epochs, and batch size, we evaluated the impact of varying the depth of the model, aiming to optimize its ability to capture the complexity of the satellite imagery.
- U-Net Configuration:
  - The U-Net model underwent a detailed testing process where learning rate, maximum epochs, number of encoders, and scale factor were all subjects of our sweeps. Special attention was given to the model's architecture adjustments, such as the number of encoders and scale factor, to refine its segmentation capabilities.

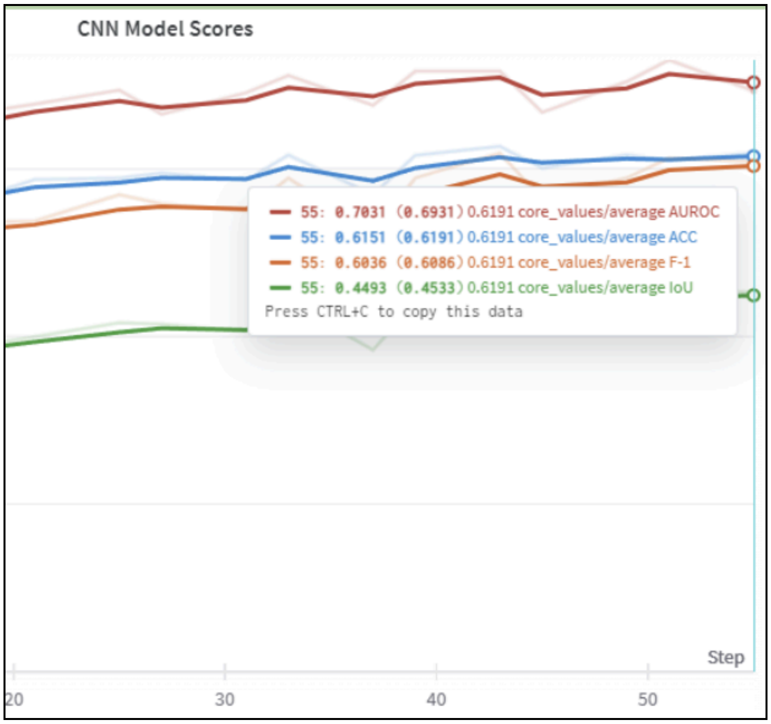
Results

Github Repository: https://github.com/cs175cv-w2024/final-project-hungry-hungry-hippo

Baseline Model Performances

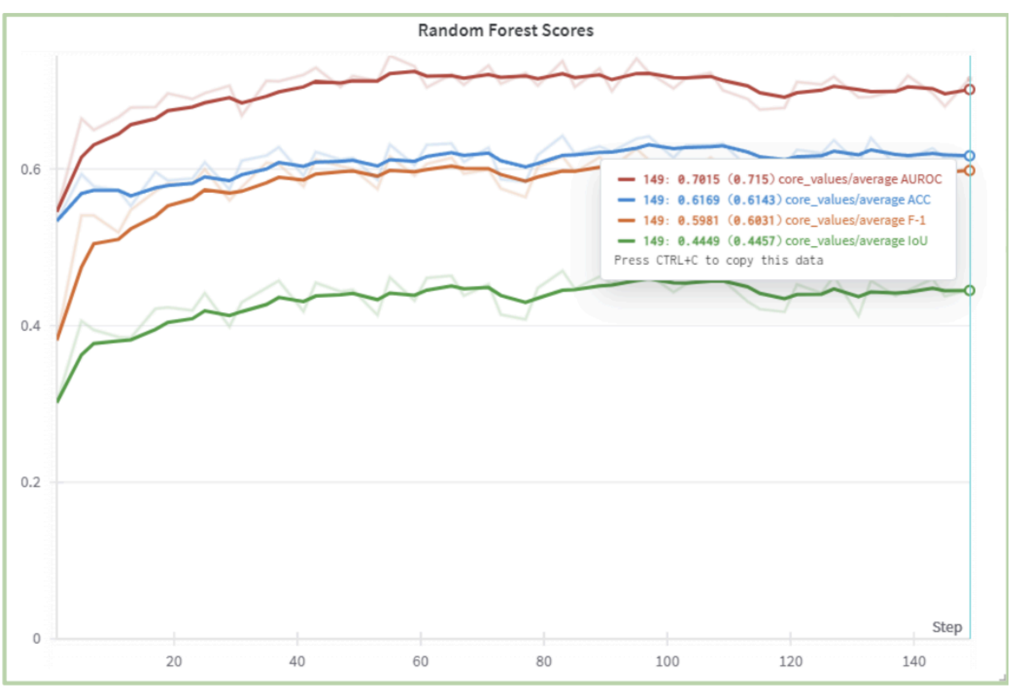






Deep Forests Classification Layer Adaptation

Following extensive tuning, training, and testing, our project reached a pivotal moment where the Segmentation CNN surpassed both ResNet and U-Net in terms of accuracy, exhibiting competitive loss and training metrics. This superiority positioned the Segmentation CNN as the primary focus for further hyperparameter optimization. Building upon this foundation, we innovated by transforming the model's output tensors into NumPy arrays, facilitating the integration of a Random Forest classifier. This adaptation actually decreased our average accuracy from 0.619 to 0.6143 average, as illustrated in the accuracy graph presented below. Spot key metrics.



CONCLUDING REMARKS

As we reflect upon the completion of this ambitious project, we are both proud and humbled by the outcomes achieved. The journey through the intricacies of satellite imagery analysis, powered by advanced deep learning techniques, has culminated in a set of remarkable results that exceeded our initial expectations. Notably, our Segmentation CNN model achieved a commendable 70% accuracy score, a testament to the efficacy of the methodologies employed and the rigorous process of parameter fine-tuning that spanned over 100 hours. The innovative adaptation of a Random Forest classifier further augmented our approach, contributing to the robustness and reliability of our findings.

The enhanced computational capabilities afforded by the NVIDIA G-Force 4060 RTX GPU played a pivotal role in our success, offering superior processing power that enabled us to delve deeper into model optimization and achieve higher accuracy scores than our peers. This technological advantage, coupled with advanced processing algorithms, was instrumental in navigating the complexities of our dataset and extracting meaningful insights from the satellite imagery.

Working on this project has been an immense honor. We extend our heartfelt gratitude to the course staff, whose expert guidance and unwavering support made this endeavor possible. Their dedication to assembling a comprehensive framework for exploration and learning provided us with the tools and knowledge necessary to tackle this challenging task.

In conclusion, this project has not only been a profound learning experience but also a demonstration of the potential that lies at the intersection of satellite technology, machine learning, and environmental analysis. The achievements documented here serve as a stepping stone towards further research and exploration in the field, with the hope of contributing to the global efforts in addressing electrification challenges. We look forward to building upon this work, inspired by the possibilities that lie ahead in harnessing technology for sustainable development and a brighter future.

REFERENCES

1. Robinson, Christopher, et al. "Large Scale High-Resolution Land Cover Mapping with Multi-Resolution Data." *arXiv*, arXiv:1805.06561, 17 May 2018, https://arxiv.org/abs/1805.06561.

2. "2021 IEEE GRSS Data Fusion Contest - Track DSE." *IEEE Geoscience and Remote Sensing Society*, https://www.grss-ieee.org/community/technical-committees/2021-ieee-grss-data-fusion-contest-track-dse/

3. "SpaceNet Challenges." *SpaceNet*, https://spacenet.ai/challenges/. Accessed [March 22, 2024].

4. Hansen, Matthew C., et al. "High-Resolution Global Maps of 21st-Century Forest Cover Change." *Science*, vol. 342, no. 6160, 2013, pp. 850-853.

5. Robinson, Christopher, et al. "Large Scale High-Resolution Land Cover Mapping with Multi-Resolution Data." *Remote Sensing of Environment*, 2019.